

subgraph2vec: Learning Distributed Representations of Rooted Sub-graphs from Large Graph

Annamalai Narayanan[†], Mahinthan Chandramohan[†], Lihui Chen[†], Yang Liu[†] and Santhoshkumar Saminathan[§]

[†]Nanyang Technological University, Singapore

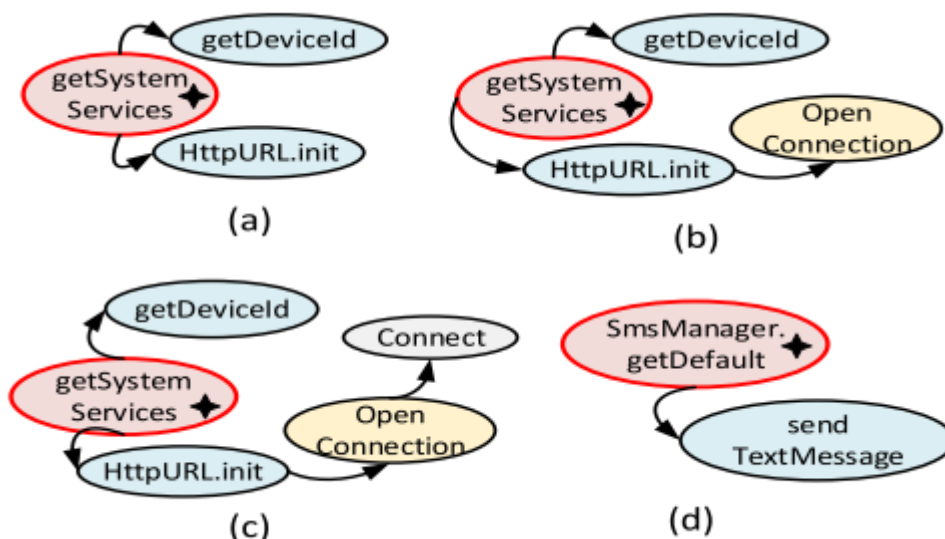
[§]BigCommerce, California, USA

annamala002@e.ntu.edu.sg, {mahinthan,elhchen,yangliu}@ntu.edu.sg, santhosh.kumar@yahoo.com

这个相当于是多了一种图元，不限定长度，这样可以抽象搜索的子结构，不拘泥于点的数量

Graph Kernels的限制

1. 结构相似



结构

相似，但是kernel不一样

2. 对角线优势 kernel的增加维度，但有些kernel图里很少，但是搜索的图有，聚类分类效果会不好

Deep Graph Kernel

过程

1. 学习表达: Learning representation of substructures
2. 上下文: Context
3. 随机游走: Deep WL Kernel

问题

1. 定义相似，有根子图，没有考虑长度，如之前的图所示
2. context 因为长度不同，某些子结构不会在同一个上下文

- 图中子结构和给定的子结构，不一定完全kernel一样，即使拓扑结构一样（拓扑算的话不一样） Every subgraph in any given graph has exactly same number of subgraphs in its context. 这是不对的
- 拓展WL relabeling strategy, 解决问题1,2
- 对skipgram model的修改，使其能够捕获不同长度的径向上下文解决问题3

贡献

- subgraph2vec 无监督学习
- 修改的skipgram model
- 讨论subgraph2vec的表示学习技术如何帮助构建WL内核的深度学习变体
- 实验

相关工作

Table 1: Representation Learning from Graphs

Solution	Learning Paradigm	node vector	subgraph vector	Context used for rep. learning
Deep Walk [8]	Semi-sup	●	○	Fixed-length random walks
node2vec [10]	Semi-sup	●	○	Fixed-Length biased random walks
Patchy-san [9]	Sup	●	●	Receptive field of sequence of neighbours of nodes
Deep Graph Kernels [7]	Unsup	●	◐	Subgraphs occurring at same degree
subgraph2vec	Unsup	●	●	Subgraphs of different degrees occurring in the same local neighbourhoods

- graph2vec 是线性的

图内核

- 有线大小图的内核: limited-size subgraphs
- 基于子树模式的内核: kernels based on subtree patterns
- 基于随意游走的内核: walk
- 基于路径的内核: path subgraph2vec是对现有内核的补充